CHERE
CENTRE FOR HEALTH ECONOMICS
RESEARCH AND EVALUATION

UNIVERSITY OF
TECHNOLOGY SYDNEY

UTS:CHERE

CENTRE FOR HEALTH ECONOMICS
RESEARCH AND EVALUATION

DERIVING WELFARE MEASURES FROM STATED PREFERENCE DISCRETE CHOICE MODELLING EXPERIMENTS

Discussion Paper

Emily Lancsar

Discussion Paper 48
August 2002

The Centre for Health Economics Research and Evaluation (CHERE) was established in 1991. CHERE is a centre of excellence in health economics and health services research. It is a joint Centre of the Faculties of Business and Nursing, Midwifery and Health at the University of Technology, Sydney, in collaboration with Central Sydney Area Health Service. It was established as a UTS Centre in February, 2002. The Centre aims to contribute to the development and application of health economics and health services research through research, teaching and policy support.

CHERE's research program encompasses both the theory and application of health economics. The main theoretical research theme pursues valuing benefits, including understanding what individuals value from health and health care, how such values should be measured, and exploring the social values attached to these benefits. The applied research focuses on economic and the appraisal of new programs or new ways of delivering and/or funding services.

CHERE's teaching includes introducing clinicians, health services managers, public health professionals and others to health economic principles. Training programs aim to develop practical skills in health economics and health services research.

Policy support is provided at all levels of the health care system by undertaking commissioned projects, through the provision of formal and informal advice as well as participation in working parties and committees.

University of Technology, Sydney
City campus, Haymarket
PO Box 123 Broadway NSW 2007
Tel +61 2 9514 4720 • Fax + 61 2 9514 4730
Email mail@chere.uts.edu.au
www.chere.uts.edu.au

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This paper was submitted as a Research Project as part of my Master of Economics in the Faculty of Economics at the University of Sydney in December 2001. I would like to thank Ms Elizabeth Savage who was my supervisor on this project, Professor Jordan Louviere who developed the experimental design used in the stated preference discrete choice modelling experiment, Professor Denzil Fiebig for advice regarding estimation of the random utility model and Associate Professor Don Wright for general comments.

ABSTRACT

The use of Stated Preference Discrete Choice Modelling (SPDCM) is gaining currency in the health economics field as a method of eliciting: preferences for goods and services; the rate at which individuals are prepared to trade off different attributes of a good or service; and the willingness to pay for goods and services. The purpose of this paper is to develop welfare measures from SPDCM data that are consistent with microeconomic welfare theory. The theory of welfare measurement using discrete data and links to the more well known literature using continuous data are presented. The estimation of welfare measures obtained from SPDCM and conjoint analysis experiments reported in the health economics literature to date are discussed, focusing on whether commonly adopted measures are consistent with microeconomic welfare theory. Finally, the Hicksian compensating variation is calculated from discrete data collected from a SPDCM experiment designed to elicit patient preferences for preventive asthma medications.

1. INTRODUCTION

The traditional approach to valuing goods and services has been to use data collected on decisions made by individuals in the market place. Such data are referred to as revealed preference (RP) data. However, RP data is often scarce and sometimes non-existent for a number of goods and services. Prominent examples exist in the health and environment sectors. In the health sector, many aspects of health care are not traded explicitly in markets, have public good characteristics, and due to universal or private health insurance, consumption is often free or heavily subsidised at the point of service.

Furthermore, due to the unique combination of characteristics of many health care markets, it may not be possible to infer consumer preferences or value from the RP data that is available[1]. Asymmetry of information between the supplier (the doctor or other health care provider) and the consumer (the patient), lack of consumer sovereignty, and the uncertain nature of both health and the outcomes of health care often result in the provider acting as an imperfect agent for the patient. This agency relationship can result in decisions (that could be collected as RP data) that are not based on the consumer's preferences alone.

One technique that could be used to overcome some of the limitations involved in using RP data to derive measures of value in the health sector is to use stated preference (SP) data, that is, the results of what individuals say they would do rather than what they are observed to do. One form of SP data collection and analysis referred to as Stated Preference Discrete Choice Modelling (SPDCM) is gaining currency within the health economics field[2]. Rather than being limited by the availability of market data, SPDCM involves the creation of a hypothetical market which can be constructed to suit the relevant research question. It can therefore be used to mimic an existing market or to elicit preferences and values for goods or services for which a market does not exist.

The results of SPDCM experiments can be used to *indirectly* elicit measures of value derived from the good or service in question in the experiment. Such measures of value have been calculated as the consumer's willingness to pay (WTP) for the good or

---

[1] The individual characteristics of many health care markets are not necessarily unique to the health care sector. What is unique however, is the combination and level of such characteristics occurring in a single market.
[2] Other forms of SP data include data collected via contingent valuation methods.

service. A small number of published studies have derived WTP measures from SPDCM experiments.

According to microeconomic welfare theory, the value of a good or services is calculated as the compensating variation (CV) or equivalent variation (EV). A key question therefore is whether the WTP measures derived from SPDCM experiments examined in the literature to date are consistent with these exact measures of value derived from microeconomic welfare theory.

In order to address this question, the theory underlying the compensating and equivalent variation is reviewed. In particular, due to the discrete nature of SPDCM experiments a method used to derive the CV and EV from discrete data, first derived by Small and Rosen (Small and Rosen 1981), is outlined. The methods used in the health economics literature to date to calculate WTP measures from SPDCM experiments are also reviewed. This review indicates that the latter are not directly consistent with the former. That is, the methods used to obtain valuations from SPDCM experiments in the literature are not directly consistent with microeconomic welfare theory.

The methods of Small and Rosen are applied in this paper to estimate the compensating variation from an SPDCM experiment designed to calculate preferences for preventive asthma medication. Despite the use of the Small and Rosen method with SPDCM data in the environment and transport economics fields with SPDCM data, it has not been used to calculate welfare measures using SPDCM data in the health economics arena to date. As such, the empirical estimation presented in this paper is the first in the health economics field to use this method with SPDCM data.

The primary purpose of this paper is to derive welfare measures from SPDCM data that are consistent with microeconomic welfare theory. To do this Section 1 provides a brief overview of SPDCM. The well known theory of welfare measurement from continuous data is reviewed in Section 2, in particular the derivation of compensating and equivalent variation. This theory is then extended to discrete data in Section 3, in particular focusing on the method of valuing a new good or service and valuing changes in the quality of an existing good or service. Section 4 reviews the estimation of welfare measures from SPDCM data in the health economics literature to date. Such measures are compared to the welfare theoretic measures outlined in Section 3. Section

5 presents an empirical estimation of an SPDCM model undertaken in a pilot study to investigate preferences for preventive asthma medication. The results of this model are used to derive the compensating variation as a measure of welfare. Section 6 presents conclusions and suggests areas for future research.

2. SPDCM

SPDCM is a method used to model consumer preferences for goods and services. Other names given to SPDCM include discrete choice experiments (DCE), discrete choice modelling (DCM), and discrete choice conjoint analysis. SPDCM assumes that individuals stated choices reflect their underlying preferences. It allows the estimation of the relative importance of the qualities or attributes that make up such goods and services and the trade offs people are prepared to make between attributes, that is, their marginal rate of substitution (MRS). It can be used to predict the probability of choosing a given commodity and market share and can also be used to estimate indirectly the compensating variation for both individual attributes and the good or service in question. Such measures of value could potentially be used in evaluation exercises such as cost benefit analyses and in resource allocation decisions.

SPDCM is based on both the hedonic principle (due to (Griliches 1961)) and Lancaster's characteristics theory of demand (Lancaster 1966): consumers have preferences for, and derive utility from, the underlying attributes rather than the commodity per se (Lancaster 1991). It is assumed that consumers purchase goods and services in order to obtain their optimal bundle of attributes and implicitly trade off these attributes when making purchasing decisions.

The technique of SPDCM involves undertaking a controlled experiment. A hypothetical market is constructed by asking respondents to choose their preferred alternative from a set of hypothetical scenarios presented in questionnaires. Each scenario is made up of a bundle of attributes that describe the good or service in question with each attribute described at one of a range of levels. For example, if the good in question was a house, the attributes and corresponding levels could be the number of rooms (1, 2, 3, 4), price ($250,000, $300,000, $350,000), whether it has a garage (yes/no) etc. The attributes remain the same in each scenario however the levels describing the attributes vary across scenarios. Respondents are thus asked to make decisions about the quality (and price) differentiated versions of the same good or service. The choices made allow analysis of which attributes are important in the decision at hand using random utility theory (RUT) (McFadden 1973).

SPDCM was first applied in transport economics and subsequently in environment economics. It has its origins in the field of marketing, however in marketing applications questions were generally framed as a ranking or rating exercise rather than a discrete choice[3]. Phrasing the question in terms of a discrete choice has two key advantages from an economic perspective. First, responses can be analysed within the economic framework of random utility theory (RUT) (McFadden 1973). According to McFadden, the Random Utility Model (RUM) is the "conceptual linkage from preferences to choices" (McFadden 1997). Second, it more accurately reflects how decisions are made in the marketplace. That is, consumers are more familiar with choosing between a discrete number of goods and services rather than ranking or rating them when making purchasing decisions (Ryan 1999).

Often, individuals do not explicitly know their preferences and they "usually cannot introspect their utility functions" (Hanemann 1996). Using SPDCM to indirectly calculate WTP values avoids one problem associated with contingent valuation studies whereby individuals are asked to explicitly state their WTP for health care. Namely, that respondents may have difficulty in explicitly stating their WTP for a good or service as they may not be accustomed to paying for health care in the market place due to the presence of universal or private health insurance. In using SPDCM, the process of having respondents make repeated choices regarding quality differentiated products encourages them to consider the trade offs they are making across attributes and allows them to reveal their preferences for the good or service in question and its underlying attributes by the hypothetical choices they make.

An advantage of SPDCM when used to measure welfare is that it does so from an ex ante perspective and as such can explicitly include any uncertainty associated with the decision which is appropriate from a theoretical point of view (Braden and Kolstad 1991). It also allows the valuation of non-use-demand which is particularly relevant in such areas as the health sector and the environment where option values can be important.

[3] In a ranking exercise respondents are asked to rank the options they are presented with from best to worst while in a rating exercise they are asked to rate an option on a scale with anchor points.

The technique of SPDCM is not without its potential limitations. A key consideration in using SPDCM is the cognitive burden it places on respondents (Bennett and Blamey 2001). Respondents are required to evaluate a number of scenarios and consider sometimes subtle differences across these in terms of the levels of the attributes and make tradeoffs in order to make decisions regarding the most preferred alternative in the choice set. The complexity of a given SPDCM experiment depends on a number of factors including: the number of scenarios; the number of attributes; the complexity of the levels of the attributes; the number of choice options in each scenario; and the respondent's familiarity with the subject area.

A related issue is that problems or inconsistencies within any of the five general steps involved in designing and estimating an SPDCM experiment (discussed in Section 6) will result in questionable results. Other potential limitations include: incentive compatibility and strategic behaviour of respondents; and framing issues or bias[4].

[4] For a discussion of each of these issues see (Bennett and Blamey 2001).

3. REVIEW OF WELFARE MEASUREMENT FROM CONTINUOUS DATA: COMPENSATING AND EQUIVALENT VARIATION

In economic theory individual welfare is measured by individual utility. In order to measure a change in an individual's welfare due to a policy change, for example, a policy that affects prices and/or income, it would therefore make sense to measure the resulting change in utility. However, due to its ordinal nature, comparing utility pre and post the policy change would only indicate the direction, but not the magnitude, of change. Nor would such an evaluation facilitate a comparison of the change in welfare across individuals, which is particularly important when a policy change makes some individuals better off and others worse off. In order to undertake such welfare analyses it is useful to have a monetary, and cardinal, measure of the change in welfare. Compensating variation (CV) and equivalent variation (EV), first introduced by Hicks (Hicks 1939)[5], offer two exact monetary measures of welfare.

The EV resulting from a proposed policy change that affects prices (or prices and income), is the change in income at *initial* prices that is equivalent in its effect on utility as the proposed price change. The EV is an ex ante measure of the change in welfare since it is measured before the price change, that is it uses the initial prices as the base. EV is negative (positive) if the price change would make the consumer worse (better) off.

In contrast, the CV is the amount of income required at the *new* prices to compensate the consumer for the change in prices thereby restoring the consumer to their initial utility level. In this sense CV is an ex post measure of the change in welfare as it is calculated using the new prices as the base. CV is negative (positive) if the price change would make the consumer worse (better) off. For a given price change, the sign of CV and EV will be the same (Varian 1992).

Thus, the EV is a measure of how much money needs to be given to or taken from a consumer *before* the price change to make them as well off as they would be if they faced the price change, that is at the new level of utility, U_1. In contrast, the CV is a

[5] Further developed by (Hicks 1942) and (Henderson 1941).

measure of how much money needs to be given to or taken from the consumer *after* the price change to leave them at their initial level of utility, U_0.

The derivation of the EV and CV starts from the Money Metric Indirect Utility Function (MMIUF) (Mas-Colell, Whinston et al. 1995); (Varian 1992), the origin of which is attributed in the literature to McKenzie (McKenzie 1957) and Samuelson (Samuelson 1974). The MMIUF is a normalisation of an ordinal utility function and is derived from the expenditure and indirect utility functions as follows:

$$MMIUF = e(\overline{P}, V(P,Y)) \qquad (3.1)$$

Where $\overline{P}$ is an arbitrary vector of reference prices, $e(.)$ is the minimised expenditure function resulting from the expenditure minimisation or "dual" problem, and $V(.)$ is the indirect utility function (IUF) derived by substituting the Marshallian demand functions into the direct utility function in the utility maximisation or "primal" problem[6]. Thus, the MMIUF indicates the minimum expenditure required to reach a given level of utility $V(P,Y)$ when prices are $\overline{P}$.[7] Using equation 2.1, the generic change in welfare (ΔW) resulting from a policy that changes prices from P_0 to P_1 can be expressed in monetary terms as:

$$\Delta W = e(\overline{P}, V(P_1,Y)) - e(\overline{P}, V(P_0,Y)) \qquad (3.2)$$

The generic ΔW can be defined as the EV or CV by choosing the vector of base prices $\overline{P}$ to be initial prices P_0 or new prices P_1 respectively. Letting $U_0 = V(P_0, Y)$ and $U_1 = V(P_1, Y)$ and noting that $e(P_0, U_0) = e(P_1, U_1) = Y$, the EV is given by:

$$\begin{aligned} EV &= e(P_{0,} U_1) - e(P_0, U_0) \\ &= e(P_{0,} U_1) - Y \qquad (3.3) \end{aligned}$$

[6] For a discussion of the primal and dual problems see (Mas-Colell, Whinston et al. 1995).
[7] The MMIUF is simply a monotonic transformation of an indirect utility function. Given any increasing monotonic transformation of an indirect utility function will produce the same preference ordering, the MMIUF is therefore itself an indirect utility function[7]. An advantage of the MMIUF is that it "involves no introspection about unobservable" (Samuelson 1974:1263).

In contrast, the CV is defined as:

$$
\begin{aligned}
CV &= e(P_{1,}U_1) - e(P_{1,}U_0) \\
&= Y - e(P_{1,}U_0) \qquad (3.4)
\end{aligned}
$$

The EV and CV can also be expressed in terms of the IUF[8]:

$$V(P_0, Y + EV) = U_1 \qquad (3.5)$$

$$V(P_1, Y - CV) = U_0 \qquad (3.6)$$

Diagrammatic representation - Indifference curves and demand curves

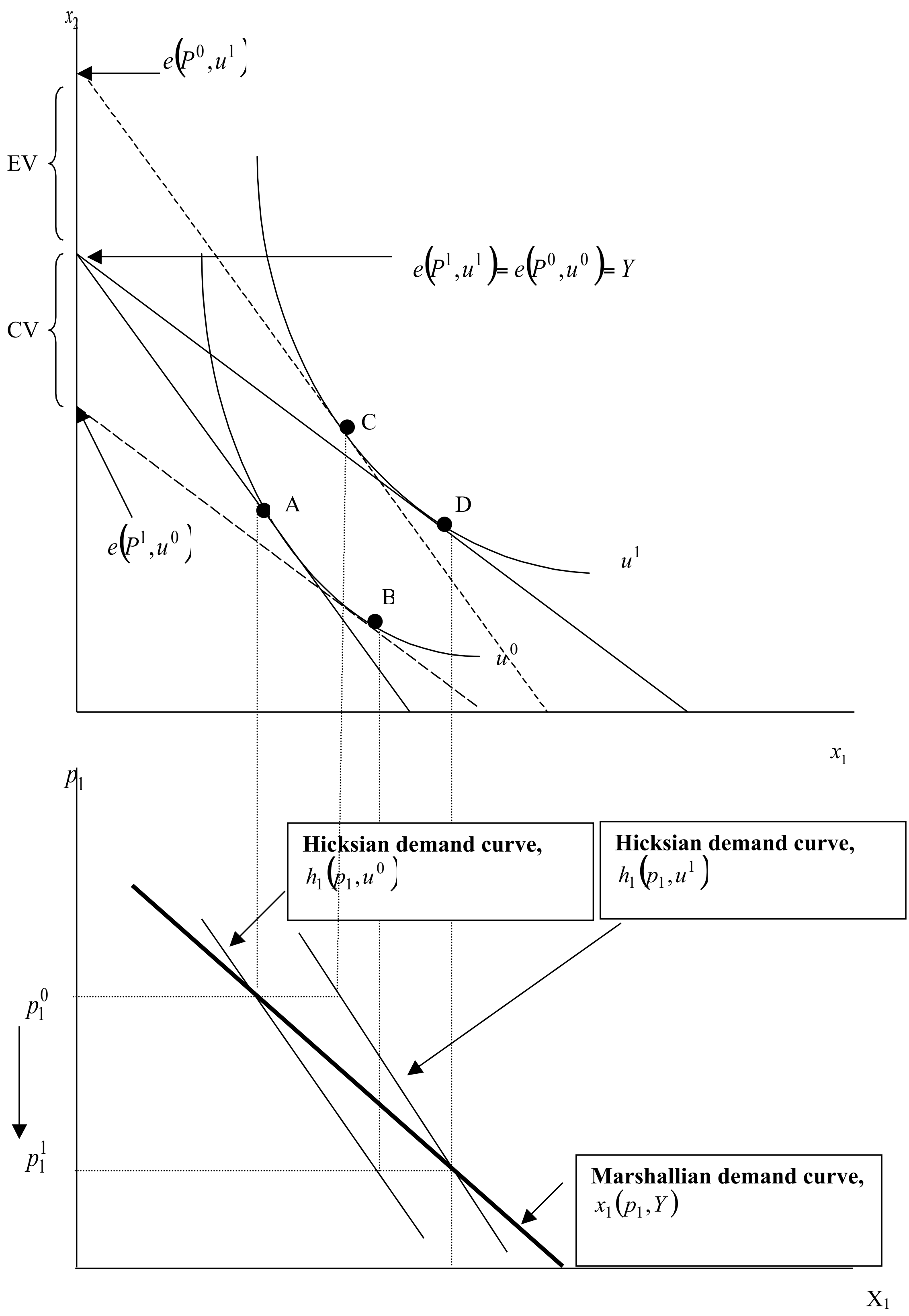
In the case of two goods the top half of Figure 1 depicts the CV and EV for a decrease in the price of good 1 holding the price of good 2 (the numeraire) constant. The consumer initially faces P_1^0, P_2^0 and is at A on indifference curve U_0. The reduction in P_1 from P_1^0 to P_1^1 reduces the relative price of P_1 which causes the budget constraint to swivel out.

The EV is the distance between the two indifference curves at initial prices. The EV indicates the increase in income required to make the consumer as well off when facing initial prices as they would be after the price decrease, that is the consumer is at point C where they face initial prices but are on the higher indifference curve U_1.

The CV is the distance between the two indifference curves at the new prices. The CV is the amount of income that would have to be taken away from the consumer at the new prices to leave them as well off a they were before the price decrease, that is at point B where the consumer faces the new prices but is on their initial indifference curve U_0.

[8] As a note of clarification, the CV and EV outlined above have been described elsewhere (King 1983) as compensating and equivalent gains. (Deaton and Meullbauer 1980) differentiate between four “consumer surplus type measures”. Two of which they denote as CV and EV which, in contrast to the measures discussed above, hold utility constant and vary prices, that is movements around a single indifference curve. The remaining two are changes in the “money metric utility” which hold prices constant and vary utility, that is movements between two indifference curves. The latter two measures (CG and EG) are the same as the CV and EV defined in the text above. With fixed income, CG = -CV and EG = - EV.

Figure 1


x_2
$e(P^0, u^1)$
EV
CV
$e(P^1, u^1) = e(P^0, u^0) = Y$
C
A
D
$e(P^1, u^0)$
u^1
B
u^0
x_1
p_1
Hicksian demand curve,
$h_1(p_1, u^0)$
Hicksian demand curve,
$h_1(p_1, u^1)$
p_1^0
p_1^1
Marshallian demand curve,
$x_1(p_1, Y)$
X_1

The EV and CV can also be represented in terms of the Hicksian, or compensated, demand curve, as depicted in the lower half of Figure 1. Using the fact that $e(P_0, U_0) = e(P_1, U_1) = Y$, the Hicksian demand curve can be recovered from the expenditure function using Shephard's Lemma as follows:

$$h_1(P,U) = \frac{\partial e(P,U)}{\partial P_1} \qquad (3.7)$$

The EV is given by:

$$EV(P_0, P_1, Y) = e(P_0, U_1) - Y$$

$$= e(P_0, U_1) - e(P_1, U_1)$$

$$= \int_{P_1^0}^{P_1^1} h_1(P_1, \overline{P}_2, U_1) dp_1 \qquad (3.8)$$

Which is represented by the area under the Hicksian demand curve at the new level of utility $h_1(P_1, U_1)$ between P_1^0 and P_1^1. Alternatively, if P_1^1 is higher than P_1^0, EV would be the negative of this area.

The representation of the CV is given by:

$$CV(P_0, P_1, Y) = Y - e(P_1, U_0)$$

$$= e(P_0, U_0) - e(P_1, U_0)$$

$$= \int_{P_1^0}^{P_1^1} h_1(P_1, \overline{P}_2, U_0) dp_1 \qquad (3.9)$$

Which is represented by the area under the Hicksian demand curve at the initial level of utility $h_1(P_1, U_0)$ between P_1^0 and P_1^1.

Magnitudes of CV and EV

In the above diagram, the EV is larger than the CV. The magnitudes of EV and CV for a price change are generally not the same which is not surprising given they are valuing the change in welfare using different prices.[9] One measures the amount of money the consumer would accept as compensation for a price change while the other measures the amount the consumer would be prepared to pay to avoid the same price change (Varian 1992). Put another way, CV and EV both measure the distance between the two indifference curves however this distance depends on the slope of the budget constraints at the point of measurement, that is it depends on relative prices. The EV is measured using initial prices while CV is measured using new prices. As such, the value of income depends on the availing prices since a dollar will purchase different amounts at different relative prices. According to Hicks, the difference depends "purely upon income effects" (Hicks 1942).

Hicks states that the magnitude of the CV and EV are the same if "the marginal utility of money is constant" (Hicks 1942). More recent treatments of this issue state that the magnitude of the CV and EV are equivalent in the restrictive case of quasilinear utility in which case, the marginal utility of income is not only constant but is equal to one and is independent of prices (Mas-Colell, Whinston et al. 1995), (Varian 1992). In this case the value of every additional dollar is independent of prices and therefore independent of the base prices used, thus there is no difference between the resulting CV and EV (Varian 1992). As a result, not only will CV be equal to EV, but they will both be equal to the Marshallian consumer surplus (CS). The CS is the area to the left of the Marshaillian demand and is therefore dependent on prices *and income*.

The relationship between EV, CV and CS

Other than in the case of quasilinear preferences, CS will be an inexact approximation to the exact change in welfare and will lie between the CV and EV. In particular, it can be seen from the Slutsky equation:

$$\frac{\partial h(P,U)}{\partial P} = \frac{\partial x(P,Y)}{\partial P} + \frac{\partial x(P,Y)}{\partial Y} x(P,Y) \qquad (3.10)$$

---

[9] This is in contrast to a change in income (rather than price), in which case the CV and EV will be the same (Hanemann 1999).

that for a normal good the derivative of the Hicksian demand curve (the term on the left hand side in (2.10)) will be larger than the derivative of the Marshallian demand curve (the second term on the right hand side in (2.10)) by the amount of the last term in equation (2.10) (Varian 1992). As such, the EV and CV will bound the CS, that is EV>CS>CV so that the CS overstates the CV and understates the EV as is the case in Figure 1. The reverse holds when the good is inferior (Mas-Colell, Whinston et al. 1995).

Willig provided upper and lower bounds on the percentage errors of approximating CV and EV with CS and suggested that if income effects are small, "the error of approximation will be very small" (Willig 1976) P589 [10]. These bounds have since been questioned in the literature (See (Hanemann 1982), (Bockstael, McConnell et al. 1991))[11]. Despite the fact that CS is an inexact measure of welfare in all but one restrictive case, the conditions of which have not been borne out by the empirical data (Slesnick 1998), it has remained popular as a method of quantifying welfare. A possible reason for its popularity is that it can be calculated directly from observed data (prices and income). The CV and EV can also be calculated from observed data but the procedure is more complex than that required to calculate CS as is discussed below.

Preferred welfare measure

Given EV and CV will produce different magnitudes of a welfare change, a valid question to ask is which is the more appropriate to use? According to Wriglesworth (Wriglesworth 1987), "the balance of opinion in the theoretical literature is in favor of the EV as the surplus measure"[12]. One advantage EV has over CV is its ability to simultaneously compare two policy changes. Since EV uses the initial prices as the base, each policy change is compared relative to the initial situation thus allowing comparison of a number of policy options to the status quo and to each other (Varian 1992). For example, EV allows the following type of statement to be made: "the consumer is x times better off under policy option 1 that moves them from (P_0, Y) to (P_1, Y) than policy option 2 that moves them to (P_2, Y). In contrast, CV cannot be used to make such simultaneous comparisons since it uses the new relative prices as the base

[10] (Hanemann 1991) showed that for a quality change the bound on the difference between CV and EV depends on both income effects and substitution effects.
[11] Chipman and Moore in particular have identified the limitations of CS (Chipman and Moore 1992)
[12] See (Wriglesworth 1987) for more detail regarding which is the best measure in terms of satisfying completeness, ordinality and cardinality.

and as the new prices will differ across policies, there will be no common base to facilitate comparison. Another advantage of EV is that "it is easier to judge the value of a dollar at current prices than at some hypothetical prices" (Varian 1992:162).

Measurement of CV and EV in practice

The CV and EV are not directly observable. Measurement of CV and EV in practice generally take one of two forms. One approach is to recover the consumer's IUF, and therefore the expenditure function since the latter is the inverse of the former, from the observable Marshallian demand functions via integration (Mas-Colell, Whinston et al. 1995). The CV and EV can then be calculated from the expenditure function. An alternative method is to specify a functional form for the IUF from which the demand functions can be derived using Roy's identity. Using the IUF and expenditure, the money metric indirect utility function and therefore CV and EV can be calculated (Varian 1992).

Relationship between EV, CV, WTP and WTA

As the majority of the studies in the health economics field that derive welfare measures from discrete data (a review of which is the focus of Section 4) refer to such welfare measures as willingness to pay (WTP) or, to a lesser extent, willingness to accept (WTA), it is useful to outline the relationship between these terms. For a reduction in price, EV measures the minimum amount the consumer is willing to accept to forgo the price change, while the CV measures the consumer's WTP to secure the change. For a price increase, the EV measures the consumer's WTP to avoid the change while the CV measures the individual's WTA compensation to tolerate the change (Hanemann 1999).

Aggregation

Aggregating welfare across individuals remains a contentious issue. One method is to simply sum the resulting individual changes in welfare and apply the potential Pareto improvement criterion which is the basis of Cost Benefit Anlaysis (CBA) (Sugden and Williams 1978). In this case, a necessary, but not sufficient, condition for a policy to improve welfare is that the sum of the change in welfare is greater than zero so that the gainers could *potentially* compensate the losers, that is the WTP of gainers is greater than the WTA compensation by the losers. Two key issues have been raised in regard to this criterion (Bockstael, McConnell et al. 1991). First, compensation is rarely paid and second, this criterion is biased in favour of the existing distribution of income since it assumes the initial distribution of income is appropriate. The latter clearly has distributional consequences. As such, estimates of WTP are dependent on ability to pay. One method to account for the disparity in income and therefore ability to pay would be to weight individual WTP to account for welfare level or position in the distribution of welfare.

4. REVIEW OF WELFARE MEASUREMENT FROM DISCRETE DATA

In many situations decisions are made between discrete or qualitative alternatives. For example, mode of transportation to work, choice of recreation site, or the choice of preventive asthma medication all involve a discrete choice from a subset of alternatives. In such situations the theory outlined in Section 3 for continuous data needs to be modified in order to estimate the changes in welfare induced from changes in either the price or quality of such goods and services. This is because the discreteness of the choices means that demand functions cannot be integrated over or alternatively indirect utility and expenditure functions cannot be differentiated (Small and Rosen 1981). As such, the conventional theory of welfare measurement must be modified to take account of discrete data. Small and Rosen completed this task in their seminal 1981 paper in which they present the theory behind the derivation of compensating variation from discrete data for both a price and quality change. The latter is discussed here[13]. Hanemann (Hanemann 1982) is also attributed in the literature with the development of methods to evaluate CV from discrete data in the field of contingent variation.

The evaluation of quality changes is of interest for many discrete choice problems. For example, a change in: the duration of a bus trip; quality of the water at a recreation site; or the required frequency of taking asthma medication could potentially result in changes in welfare for the individuals considering the choice of mode of transport, recreation site or asthma medication, respectively.

Derivation of CV for a quality change

The following derivation of the CV for a quality change follows that of Small and Rosen (Small and Rosen 1981). Consider the case in which there are three commodities that are available in continuous quantities, x_1, x_2 and x_n where x_n is the numeraire good. Goods x_1 and x_2 are mutually exclusive, hence:

$$x_1 x_2 = 0 \qquad (4.1)$$

[13] See Small and Rosen (1981) for the derivation of the CV for a price change using discrete data. Alternatively, price can be treated as an attribute of the good or service and the CV resulting from a price change can be calculated in the same way as for a quality change.

which prevents both x_1 and x_2 being consumed in positive amounts. The discrete choice is therefore between x_1 and x_2. One of the commodities, say good 1, has a quality variable associated with it. The twice differentiable, strictly quasi-concave utility function is:

$$U = U(x_n, x_1, x_2, q_1) \qquad (4.2)$$

where the scalar q_1 describes the quality variable which is assumed to be exogenous to consumers. Utility is assumed to be finite whenever either x_1 or x_2 are zero. It is also assumed that the quality of good 1 is irrelevant unless good 1 is being consumed, that is:

$$\frac{\partial U(x_n, 0, x_2, q_1)}{\partial q_1} = 0 \qquad (4.3)$$

Consumers maximise 4.2 subject to the budget constraint and non negativity constraint:

$$x_n + p_1 x_1 + p_2 x_2 = y \qquad (4.4)$$

$$x_j \geq 0 \quad (j = 1,2,\ldots,n) \qquad (4.5)$$

An interior solution is ensured by assuming that the numeraire and either x_1 or x_2 are consumed in positive amounts.

Let $U_0 = V(p_1^0, p_2, q_1, y)$ be the value of the IUF at initial prices and income, $U_1 = V(p_1^1, p_2, q_1, y)$ be the IUF at final prices, and $e = (p_1, p_2, U)$ be the minimum expenditure required to achieve utility U. The IUF and expenditure function satisfy:

$$U = V(p_1, p_2, q_1, e(p_1, p_2, q_1, U)) \qquad (4.6)$$

Taking the quality derivative of equation (4.6) produces:

$$\frac{\partial U}{\partial q_1} = \frac{\partial V}{\partial q_1} + \frac{\partial V}{\partial e}\frac{\partial e}{\partial q_1} \qquad (4.7)$$

Setting $\frac{\partial U}{\partial q_1} = 0$ and noting that e is equal to y and rearranging yields:

$$\frac{\partial e}{\partial q_1} = -\left(\frac{1}{\lambda}\right)\frac{\partial V}{\partial q_1} \qquad (4.8)$$

Where λ denotes the marginal utility of income which is equal to $\partial U / \partial x_n$. The marginal utility of income is used in equation (4.8) to convert the marginal utility of quality into monetary units. Constant marginal utility of income is often a helpful, but not a necessary, condition for equation (4.8) to hold.

The change in the area under the compensated demand curve after it has shifted in response to a change in quality is presented in Small and Rosen's Theorem 2 as

$$\frac{\partial}{\partial q_1} e\left(p_1^0, p_2, q_1, U_0\right) = -\int_{p_1^0}^{\infty} \frac{\partial}{\partial q_1} h_1\left(p_1, p_2, q_1, U_0\right)\, dp_1 \qquad (4.9)$$

For a change in quality from q_0 to q_1, integrating over 4.9 produces the CV

$$\Delta e = -\int_{p_1^0}^{\infty} \left[h_1\left(p_1, p_2, q_1^1, U_0\right) - h_1\left(p_1, p_2, q_1^0, U_0\right)\right]\, dp_1 \qquad (4.10)$$

While this is a useful result if data are available to estimate the compensated demand curve which can be integrated over, stated preference models do not usually generate demand curves. Fortunately, the CV can be derived directly from the expenditure function and from random utility models which do not depend on demand functions (Bockstael et al 1991).

Application to Random Utility Theory

Given the inherent uncertainty regarding which alternative will be chosen in SPDCM experiments, such models are estimated using random utility theory (RUT) and econometric models. As the ultimate aim of this paper is to derive welfare measures from SPDCM experiments, it is necessary to investigate how the CV formula outlined above is calculated from stochastic utility models.

RUT is based on the premise that although the individual knows their own preferences with certainty, some components of their preferences are unobservable to the researcher, and therefore treated as random (McFadden 1973). The IUF, for individual i conditional on choice of good j, U_{ij}, is assumed to be additively separable into 2 components:

$$U_{ij}(p_i, q_i, y_i) = V_{ij}(p_j, q_j, y_i; Z_i) + \varepsilon_{ij} \qquad (4.11)$$

The first term, V_j, is the deterministic or non stochastic component, the *form* of which is assumed to be the same for all individuals in the population. This term is a function of price, quality attributes, income and the vector Z_i, which contains observable characteristics of the individual. The final term is the realisation for individual i of the unexplainable or stochastic component and is assumed to be independent of the components of V_{ij}.

The systematic part of the IUF can be described as follows:

$$V_{ij} = \beta' X_{ij} + \gamma' Z_i \qquad (4.12)$$

where X_{ij} is the vector of attributes, including price and quality, of the jth good as viewed by the ith individual and Z_i is a vector of personal characteristics for individual i including income.

The IUF represents the maximum attainable utility for given prices and income. There are four properties of an IUF. First, it must be homogeneous of degree zero in prices and income. This means that multiplying both prices and income by a positive constant will have no effect on the IUF. Second, the IUF must be non increasing in prices. This implies that as price rises utility must not increase. Third it must be quasiconvex and finally it must be continuous. The first two imply constraints on the IUF which is estimated in SPDCM experiments. One way to impose the property of homogeneity of degree zero is to enter price and income as a ratio in the model to be estimated. The implication of the second property that the IUF be non increasing in price is that the coefficient on the price attribute in the random utility model should be non positive.

In the RUT framework an individual chooses the alternative from the choice set that results in the highest utility (McFadden 1973). Thus, selection of one alternative implies that the utility derived from that alternative is greater than the utility associated with all other options in the choice set.

Due to the fact that overall utility is random, only the probability of choice of one option over another can be analysed (Adamowicz, Boxall et al. 1999). Individual i will choose good 1 if and only if the utility derived from good 1 is greater than the utility derived any other good in the choice set J. Assuming a joint probability distribution for ε_i, the probability that utility is maximised by choosing good 1 is given by:

$$\begin{aligned}\pi_{i1} &= pr(U_{i1} > U_{ij}) \\ &= pr(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij}) \\ &= pr(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}) \quad \forall j \neq 1 \qquad (4.13)\end{aligned}$$

Where π_{i1} is conditioned on p, q and Z. It is the prediction of the fraction of individuals who choose good 1. Thus, the probability that good 1 will be chosen depends on whether the utility derived from option 1 is greater than the utility derived from all other options.

Assuming a Type 1 extreme value error distribution with scale parameter μ produces the logit specification of the choice probability. This results in the following expression:

$$\pi_{i1} = \frac{e^{\mu V_{i1}}}{\sum_{j=1}^{n} e^{\mu V_{ij}}}, \quad j = 1,\ldots,n \qquad (4.14)$$

where the scale parameter μ is usually set equal to one (Adamowicz, Louviere et al. 1994). The scale parameter cannot be identified in a single sample but can be identified if two separate samples are analysed together. The relative scale parameter accounts for the difference in the variation of the unobserved effects (Adamowicz et al 1994, Swait and Louviere 1993).

Using equation (4.12), equation (4.14) can be rewritten as:

$$\pi_{i1} = \frac{e^{\beta' X_{i1} + \gamma' Z_i}}{\sum_{j=1}^{n} e^{\beta' X_{ij} + \gamma' Z_i}}, \quad j = 1,...,n \qquad (4.15)$$

The CV is given by:

$$CV = -\frac{1}{\lambda} \int_{V_1^0}^{V_1^1} \pi_1 (V_1, V_2) dV_1 \qquad (4.16)$$

Given the logit specification of the choice probability, the integral in equation (4.16) can be explicitly evaluated as[14]:

$$CV = \frac{1}{\lambda} \left[\ln \sum_{j=1}^{n} e^{V_j^0} - \ln \sum_{j=1}^{n} e^{V_j^1} \right] \qquad (4.17)$$

This is equivalent to the formula proposed by (McFadden 1999). The term $\ln \sum_{j=1}^{n} e^{V_j}$ has been referred to in the environmental and transport literature as the inclusive value (Bockstael, McConnell et al. 1991) and (Ben - Akiva and Lerman 1985). Ben - Akiva and Lerman define the inclusive value as “the expected maximum utility of a subset of alternatives” (Ben - Akiva and Lerman 1985: 282). The inclusive value weights the IUFs associated with each alternative by the probability of choosing each alternative (Bockstael, McConnell et al. 1991). As such, the CV formula is related to the probability of choosing a given alternative: a low probability of choosing a given option results in a low CV while a high probability of choosing a given option results in a high CV (Bockstael, McConnell et al. 1991). This relationship between the CV and the probability of choosing each alternative is not surprising given the derivative of the term in the square brackets in (4.17) with respect to V is equation (4.16). In order to

[14] Alternatively a probit model could be estimated by assuming a standard normal error distribution rather than a logistic distribution. For the derivation of the CV formula for a probit model see Small and Rosen (1981).

express the change in utility in monetary terms, the difference in the inclusive values in equation (4.17) is scaled by λ, the marginal utility of income (Trajtenberg 1989). The marginal utility of income is given by the coefficient on income in the estimated IUF. However, often information on income is not available. This has been addressed in a number of empirical applications by interpreting the coefficient on the price attribute (which represents the marginal disutility of price) as the negative of the marginal utility of income (Louviere, Hensher et al. 2000). In this case the CV formula is represented by:

$$CV = -\frac{1}{\beta_p}\left[IV_0 - IV_1\right] \quad (4.18)$$

where β_p is the coefficient on the price variable.

Although equation (4.17) is referred to as the CV, it is in fact equivalent to the EV and Marshallian CS if the marginal utility of income (represented by λ) is assumed to be constant and equal to one (see Section 3).

In practice, equation (4.17) can be used to derive welfare measures arising from a change in a single quality attribute in a single good or a change in all quality attributes associated with a good. It can also be used to value the benefits associated with the introduction of a new good or service if its characteristics can be described (Bockstael, McConnell et al. 1991). In such cases the CV is given by:

$$CV = \frac{1}{\lambda}\left[\ln\sum_{j=1}^{n+1} e^{V_j} - \ln\sum_{j=1}^{n} e^{V_j}\right] \quad (4.19)$$

where V_{n+1} is the systematic component of the IUF of the new good or service. Alternatively, the losses due to the elimination of a good or service can be calculated as:

$$CV = \frac{1}{\lambda}\left[\ln\sum_{j=1}^{n} e^{V_j} - \ln\sum_{j=2}^{n} e^{V_j}\right] \quad (4.20)$$

(Bockstael, McConnell et al. 1991)

Rate of change in the CV as quality improves

Since the unobserved factors are additive in equation (4.11), the rate of change in the CV for an infinitesimal change in quality can be calculated (McFadden 1999). For example, if quality improves infinitesimally from q_0 to $q_1 = (q_0 + \Delta q)$ which results in an infinitesimal change in the choice probabilities $\pi_{1j} = (\pi_{ij} + \Delta\pi)$ then the compensating income reduction for a fixed alternative j is given by:

$$CV = \left[\frac{\partial V / \partial x_j}{\partial V / \partial y}\right]\Delta x_j + \sigma(\Delta x) \qquad (4.19)$$

where $\sigma(\Delta x)$ is infinitesimal and can therefore be ignored for practical purposes. The use of this formula is appropriate if the choice option of interest is chosen with certainty or if the change in quality is small (McFadden 1999).

Aggregation

The convention in aggregating welfare measures derived from discrete data over individuals has been to multiply the average CV of the sample by the population size (Bockstael, McConnell et al. 1991). The same procedure can be carried out for segments of the population based on location or income strata, for example. Following the potential Pareto improvement criterion, if the benefits net of costs are greater than zero the gainers could compensate the losers thus implying that the change (potentially) results in a net benefit.

Empirical Applications

Equation (4.17) has been used in a number of empirical applications in the areas of environmental and transport economics to calculate the CV from discrete data.
It is important to point out that this formula has been used in environment and transport economics to derive welfare measures not only from SPDCM data such as (Adamowicz, Swait et al. 1997) which investigated the inclusion of and difference between perceptions and objective data, but also from a number of other forms of discrete data. For example, discrete RP data (Trajtenberg 1989), referendum contingent valuation data (Adamowicz, Boxall et al. 1999), combined RP and SP data (Adamowicz, Louviere et al. 1994; Adamowicz, Swait et al. 1997; Louviere, Hensher et

al. 2000) and combined contingent valuation and SPDCM data (Adamowicz, Boxall et al. 1999).

Despite its use in the environmental and transport economics fields, as will be discussed in Section 5, this method has not been used explicitly with SPDCM data to elicit welfare measures in the health economics arena to date. However, it has been used with RP data in the health economics field. For example, this method has been used by (Gertler 1987) to investigate the welfare implications of user fees in Peru, by (Feldman 1994) in an analysis of the change in consumer surplus from a merger of health plans, and by (Puig-Junoy 1998) to derive the compensating variation for emergency visits, while the probit form has been used by (Clarke 1998) in a travel cost approach applied to mammographic screening.

5. REVIEW OF WELFARE MEASUREMENT FROM SPDCM DATA IN THE HEALTH ECONOMICS LITERATURE

SPDCM has increasingly been used in the health economics field to model preferences for health care related goods and services. In doing so, a number of studies have used the coefficients from the estimated IUFs to calculate measures of willingness to pay for individual attributes and for entire goods and services. This section provides an audit of those studies which have calculated WTP from SPDCM data, briefly discusses the methods used to derive WTP values from SPDCM data in the literature and compares these to the theoretically correct method set out in Section 3.

A review of the health economics literature was undertaken. The search strategy is outlined in Appendix One. The review found 10 studies had calculated the WTP for a change in a single attribute, referred to in the literature as the "marginal WTP", and 7 studies that had estimated the WTP for a an entire good or service from SDPCM or conjoint analysis data. These two groups are not mutually exclusive.

WTP for individual attributes

Recall that the IUF can be described as follows:

$$V_{ij} = \beta' X_{ij} + \gamma' Z_i \qquad (5.1)$$

where X_{ij} is the vector of attributes, including price and quality, of the jth good as viewed by the ith individual and Z_i is a vector of personal characteristics for individual i.

Upon estimating the coefficients on the attributes, the β s, in equation (5.1), a number of studies[15] have interpreted the marginal rate of substitution (MRS) between an individual attribute and the price attribute as the marginal willingness to pay for a change in a single attribute as follows:

[15] (Ryan and Huges 1997; Ryan 1999; Ryan 1999; Gyrd-Hansen 2000; Ryan 2000; Scott 2001)

$$MWTP = -\frac{\partial V / \partial X_1}{\partial V / \partial P} = \frac{\beta_1}{\beta_P} \qquad (5.2)$$

where X_1 represents attribute 1, the attribute of interest, and P is price. The numerator is interpreted as the marginal utility of this attribute and the denominator is the marginal disutility of price or cost.

This interpretation is appropriate if the alternative whose attribute has changed is chosen with certainty. However, which alternative will be chosen in a stated preference discrete choice experiment is inherently uncertain. Interpreting the ratio of the coefficient on the attribute of interest to the price coefficient as the marginal welfare effect is "not entirely consistent with the random utility model if the welfare effect being examined is a change in one of several possible alternatives", in which case the probability of choosing that alternative must be taken into account (Adamowicz, Boxall et al. 1999:467). In contrast, the equation outlined in Section 3 for calculating the CV (equation 4.17) implicitly weights the welfare change by the probability of choice and is therefore consistent with random utility theory. McFadden suggests that equation (5.2) will approximate WTP for small quality improvements (McFadden 1999).

WTP for goods or services

A number of studies[16] have calculated the WTP arising from a change in all levels of a good or service (which is implicitly the same as estimating the WTP for a new entire new commodity) using the following formula:

$$WTP = \sum_k \frac{\beta_k}{-\beta_p}(\Delta X_k) \qquad (5.3)$$

where subscript k represents the attributes describing the good or service in question and subscript p denotes price. This method simply involves the summation of the product of marginal willingness to pay described above in equation (5.2) and the change in levels across all attributes. As with the case of the marginal willingness to pay formula, this formula is not consistent with random utility theory as it does not take

[16] (Ryan and Huges 1997; Johnson, Desvousges et al. 1998; Ryan 1999; Gyrd-Hansen 2000; Johnson, Mathews et al. 2000; Ryan 2000)

account of the probability of choosing the good in question and is appropriate only if that alternative is chosen with certainty. If, however, the good or service is one of a number of alternatives in the choice set, as is the case in SPDCM experiments, then the probability of choosing that alternative should be taken into account. As outlined in Section 4, a change in a good or service with a low probability of being chosen will have a small welfare impact while the converse case will have a large welfare impact (Bockstael, McConnell et al. 1991). For example, a change in the quality attributes of a good in the choice set which has a very low probability of being chosen should result in a low WTP since changing an attribute of a good which no one chooses will have little effect on welfare.

An important issue is whether the method used in the health economics literature to date to derive WTP values from SPDCM data (equation 5.3) is ever consistent with Small and Rosen's method (equation 4.17). The answer is yes, but only under two very specific assumptions. First that there are only two choice options in the choice set and second if an approximation is made. To show the circumstances in which the two equations are equivalent, equation 5.3 is first rewritten as:

$$WTP = \frac{1}{-\beta_p} \sum_k \beta_k \left(\Delta X_k \right) \qquad (5.4)$$

Since X_k represents the difference in the levels of the attributes across the two alternatives in the choice set, equation 5.4 can be simplified further to:

$$WTP = \frac{1}{-\beta_p} \sum_k \beta_k \left(X_k^0 - X_k^1 \right) \qquad (5.5)$$

$$WTP = \frac{1}{-\beta_p} \left(V^0 - V^1 \right) \qquad (5.6)$$

For equation 5.6 to be equivalent to Small and Rosen's formula (equation 4.17) two requirements must be met. First, the choice set must contain only two alternatives, that is a binary choice, that is if the choice set contains 3 or more alternatives 5.6 is not equivalent to 4.17, and second the following approximation must be imposed:

$$\ln\left(1 + e^{v_0}\right) \approx V_0$$

and

$$\ln\left(1 + e^{v_1}\right) \approx V_1 \qquad (5.7)$$

In this case equation 4.17

$$CV = \frac{1}{\lambda}\left[\ln \sum_{j=1}^{n} e^{V_j^0} - \ln \sum_{j=1}^{n} e^{V_j^1}\right]$$

collapses to:

$$CV = \frac{1}{\lambda}\left[\ln\left(1 + e^{V_j^0}\right) - \ln\left(1 + e^{V_j^1}\right)\right] \quad (5.8)$$

$$CV = \frac{1}{\lambda}\left(V^0 - V^1\right) \quad (5.9)$$

which is equivalent to equation (5.6) where $-\beta_p$ is interpreted as the marginal utility of income λ. In general, however, the method used in the health economics literature to date (equation 5.3) is not equivalent to the theoretically consistent method (equation 4.17).

6. EMPIRICAL APPLICATION: DERIVING WELFARE MEASURES FROM AN SPDCM EXPERIMENT

Having outlined the methods of deriving welfare measures from discrete data that are consistent with microeconomic welfare theory, this section presents an empirical application that aims to put that theory into practice. In particular, a SPDCM experiment is developed and estimated using data on patient preferences for preventive asthma medication[17]. The coefficients from this estimated model were then used to calculate the compensating variation.

Unlike many other survey instruments, a new SPDCM questionnaire is developed for each new choice experiment. There are five main steps involved in developing and estimating an SPDCM experiment: determination of attributes and levels; experimental design; questionnaire design; data collection; and analysis. Each step is outlined below with reference to the asthma SPDCM experiment.

1. Attributes and levels

The first stage in developing an SPDCM experiment is to identify a number of attributes or characteristics that describe the decision at hand. The attributes and levels associated with the choice of asthma medication were identified from: the literature; discussions with respiratory clinicians; and most importantly from structured interviews with asthmatics. The cost attributes were chosen to represent the spread of prices of preventive asthma medication. $0 is the lower bound and $100 is the upper bound. $20 was selected to be close to the Pharmaceutical Benefits Schedule (PBS) co-payment price of $21.90. $80 was chosen to represent the cost of asthma medications not listed on the PBS. Allowing for a large spread in the cost levels is important in order to ensure that the maximum WTP is elicited. The asthma SPDCM experiment involved 10 attributes, 7 of which have 4 levels, the remaining 3 have 2 levels. The attributes and levels are contained at Table 1.

In health economics goods and services have typically been valued based on health outcomes, for example, number of deaths prevented, reduction in blood pressure etc.

[17] The asthma SPDCM study discussed here was a pilot study of a questionnaire that is part of a 3 period cross over randomised controlled trial of three preventive asthma medications being conducted by the Cooperative Research Centre for Asthma.

SPDCM recognises that in addition to health outcomes, non health outcomes and process attributes may also be important in decision making regarding health care (Ryan 1999). As such, the attributes identified as important in the choice between preventive asthma medications have been sorted into these three sub groups in Table 1. The measures of value derived from this SPDCM experiment will therefore include more than health outcomes.

Table 1: Attributes and Levels

Attribute	Description	Levels
Process Attributes		
Cost	Total cost of the drug to you for a (3) month supply is:	• $0 • $20 • $80 • $100
Repeats	In order to pick up your prescription and repeats for each 6 month period you will have to go to the chemist:	• once • twice • 3 times • 6 times
Administration	You take the drug via:	• an aerosol inhaler • a dry powder inhaler, e.g. Turbuhaler or Accuahaler • a tablet • both an inhaler and a tablet
How often	You need to take the drug:	• once a day • twice a day
Monitoring	You measure your morning peak flow:	• every morning • never
Health Outcomes		
Symptom Severity	On this drug you will experience:	• minimal symptoms once a week or less • exercise breathlessness, cough or wheeze once a week • chest tightness doing normal activities requiring reliever twice a week • awaking at night with asthma requiring the use of reliever more than 4 times a week
Usual daily activities	Compared to most people in your age group you are able to participate in:	• all usual daily activities without the use of a reliever • all usual daily activities provided you use a reliever for some of these activities • most usual daily activities provided you use your reliever most of the time • a few usual daily activities provided you use a reliever all the time
Sporting/strenuous activities	You are able to participate in:	• all the sporting or strenuous activity you want without difficulty • all the sporting or strenuous activity you want with the use of reliever • a restricted range of the sporting and strenuous activity you want with the use of a reliever • no sporting or other strenuous activity
Side effects	On this drug you will experience:	• no side effects • tremors, palpitations, nervousness or headache • oral thrush • occasional hoarseness of speech
Non Health Outcomes		
Dr recommendation	Your doctor says that this drug:	• is the best for your asthma • will give you satisfactory control of your asthma

2. Experimental design

The experimental design was developed by combining the levels of the attributes into scenarios which respondents saw in the questionnaire[18]. Given the experiment involves 7 attributes with 4 levels and 3 attributes with 2 levels, the full factorial design consists of $4^7 \times 2^3$ = 13,1072 scenarios, clearly too large to be completed by a single respondent. To reduce the number of scenarios, an orthogonal fractional factorial design was employed which contained 256 scenarios. Two designs were originally combined in this study. The first design contained 12 scenarios using the extreme levels, that is the first or last level on each attribute. The second design contained 256 scenarios and was broken into 16 random versions of 16 scenarios. The rationale for combining two designs is it allows all respondents to face an identical set of 12 scenarios which allow for greater statistical precision when investigating the end points of the design (the lowest and highest levels on each attribute) and also allows the respondent population to be segmented into different groups. The design consisting of 16 random versions of 16 scenarios allows investigation of the response surface.

A number of scenarios were altered in the design due to implausible combinations of three attributes: the daily activities, symptoms and sporting activities. These combinations were randomly replaced with plausible combinations. An example of an implausible scenario was one that stated that on the hypothetical asthma medication the individual would be able to participate in “all the sporting or strenuous activity you want without difficulty”, and yet have the worst symptom level “waking at night with asthma requiring the use of reliever more than 4 times a week” and the worst daily activity level, namely that the individual is only able to participate in “a few usual daily activities provided [they] use a reliever all the time”. The resulting design was still reasonably orthogonal and balanced in terms of the number of times each level of an attribute was seen in a scenario. Thus, a trade off was made between full orthogonality and the plausibility of the scenarios since there is the possibility that respondents will not take the questionnaire seriously if the scenarios are unrealistic.

[18] The experimental design used in this pilot study was developed by Professor Jordan Louviere.

3. Questionnaire Design

The third step involved the development of the questionnaire in which respondents were presented with the hypothetical scenarios. The design containing 12 scenarios was combined with a single random version of 16 scenarios to produce a questionnaire containing 28 scenarios. The questionnaire consisted of a “report card” in which respondents ranked their current asthma medication against the same attributes and levels used in the design; demographic questions and 28 hypothetical scenarios. Under each scenario respondents were faced with 3 choices: the hypothetical asthma medication outlined in the scenario; their current asthma medication as described in the report card; or no asthma medication. The third option, no asthma medication, was included to ensure respondents were not forced to make a choice between two alternatives, of which they may choose neither in practice. An example of a scenario and choice question are contained in Appendix 2.

4. Data Collection

The pilot study was conducted in Sydney, Australia, in August 2000 on a sample of 30 asthmatics obtained from an existing database. Three interviewers were used in this study, each with a sub sample of 10 respondents. On completion of the questionnaire all respondents were asked specific and open ended debriefing questions regarding the questionnaire. In particular, concerning: the ease of answering; the relevance of attributes and appropriateness of levels; and how respondents chose between the 3 options they were presented with; and whether there were some attributes that were more important than others in making this decision.

5. Model estimation

The choice between the three asthma medication options was modelled in a random utility framework. As outlined in Section 4, RUT proposes that utility, which is a latent variable, can be separated into an explainable and a stochastic component as follows:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \qquad (6.1)$$

where U_{ij} is the utility derived from choice j by individual i, V_{ij} is the explainable or systematic component of utility and ε_{ij} is the random component. Assuming a linear functional form, the systematic component of the IUF can be described as follows:

$$V_{ij} = \beta' X_{ij} + \gamma' Z_i \quad (6.2)$$

where X_{ij} is the vector of attributes of the jth asthma medication choice as viewed by the ith individual and Z_i is a vector of personal characteristics for individual i (Maddala 1983). Thus β and γ are vectors of parameters which represent the influence of the medication attributes and individual characteristics on the systematic component of the IUF. In this model the effect of each attribute, β, on the IUF is assumed to be homogeneous across all respondents.

Respondents choose the alternative j to maximise their utility given the alternatives in the choice set. Individual choice behaviour is assumed to be probabilistic. The probability that a respondent will choose a given alternative is represented as:

$$\begin{aligned} \pi_{i1} &= \Pr(Y_i = 1 \mid J) \\ &= \Pr(U_{i1} > U_{ij}) \\ &= \Pr(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij}) \\ &= \Pr\left(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}\right) \quad \forall j \neq 1 \end{aligned} \quad (6.3)$$

where π_{i1} is the probability individual i will choose asthma medication option 1 given the choice set J.

Assuming a Type 1 extreme value error distribution produces the conditional logit specification of the probability of choice:

$$\pi_1 = \frac{e^{V_1}}{\sum_{j \in C} e^{V_j}} \quad (6.4)$$

The logit specification implies a difference in utility model. As such, the effect of the individual characteristics, which by definition are the same for a given individual across the j choice alternatives, will cancel out in the estimation process. In order to include the effect of individual characteristics, they must be interacted with a variable that does vary, such as one of the attributes or a dummy variable created for the choice alternatives, also referred to as alternative specific constants. In this study individual characteristics such as income, age and gender were interacted with the alternative specific constant.

Values of the two attributes in the design that can be interpreted as numerical or ordered variables, cost and number of repeats, have been coded by centring them around their mean. For example, the mean of the levels on the cost attribute ($0, $20, $80, $100) is $50 which was then subtracted from each of the levels to give -$50, -$30, $30 and $50. As such, the impact of the mean enters the constant. The remaining eight attributes are qualitative and do not have a clear ordering. These variables have been effects coded. An example of effects coding for a four level attribute is contained in Table 2. Parameters are estimated for the first 3 effects coded levels while the parameter on the last level of each attribute is obtained by taking the negative sum of the estimated parameters on the first 3 levels. An advantage of effects coding rather than dummy coding is that the former does not confound the effect of the attributes with the alternative specific constant while the latter does. Effects coding orthogonalises the effect of the attributes to the constant (Adamowicz, Louviere et al. 1994).

Table 2: Effects coding for a four level attribute

Levels	Fx1	Fx2	Fx3
0	1	0	0
1	0	1	0
2	0	0	1
3	-1	-1	-1

Results

The SP responses to the choice of one of the 3 asthma medication alternatives: hypothetical medication; current medication; or no medication, were analysed using a multinomial logit model. As with all conditional logit models, one of the sets of parameters must be normalised to equal zero and used as the reference to which the estimates for the other two alternatives are compared. In this case the coefficients for the "no medication" option were set to zero which implies that the utility of choosing the no medication option is also set at zero[19].

[19] Choice of the no medication alternative as the reference was based on the fact that the database consisted of information on the levels of the attributes of the hypothetical scenarios and the respondents current medication (collected via the report card) but did not contain information regarding the levels of the attributes associated with not taking an asthma medication, so less information was lost.

The results indicated that the majority of respondents chose to remain on their current asthma medication and that the probability of choosing the "no medication" option was very small, 0.0093 or less than one percent. Given the very low number of times the "no medication" option was chosen (21 out of 360 observations) and due to the small numbers used in this pilot study (n=30), use of the no medication option as the reference would result in statistical imprecision of the estimated parameters. As a result, this choice option was dropped from the data set as were those 3 respondents who chose this alternative[20]. This left 27 respondents with 324 observations. The model was then estimated in STATA as a binary conditional logit model where the relevant choice was between the hypothetical medication and the respondent's current medication. The results of this model are presented in Table 3.

Table 3: Results of Asthma Medication Discrete Choice Model

Attributes	Coefficient[a]	Standard Error	P value
Constant	-3.820	0.738	0.000*
Cost	-0.015	0.002	0.000*
Repeat	-0.033	0.082	0.684
Adminfx1	-0.610	0.365	0.095***
Adminfx2	-0.139	0.506	0.783
Adminfx3	0.780	0.771	0.312
Oftenfx1	0.326	0.207	0.116
Sympfx1	1.100	0.374	0.003*
Sympfx2	0.758	0.414	0.067***
Sympfx3	-1.588	0.569	0.005*
Activfx1	0.169	0.260	0.516
Activfx2	0.024	0.408	0.952
Sportfx1	0.808	0.363	0.026**
Sportfx2	1.674	0.566	0.003*
Sportfx3	-1.645	0.553	0.003*
Monitfx1	0.024	0.181	0.896
Sidefx1	0.792	0.467	0.090***
Sidefx2	-1.030	0.641	0.108
Sidefx3	0.903	0.921	0.326
Number of observations		634	
Pseudo R^2		0.4912	
LR $Chi^2(19)$		382.29	

Notes:
(a) * significant at 1% level, ** significant at 5% level, *** significant at 10% level
(b) LR test for null hypothesis that all coefficients except the intercept are jointly zero; critical value at 1% level is approximately 36.1908.

[20] Eliminating choice options due to the low number of respondents who chose that option has been undertaken elsewhere. See, for example, (Louviere, Hensher et al. 2000) in which case one of the four

The model provided a good fit to the data with a pseudo R^2 of 0.4912. The likelihood ratio test for the null hypothesis that all coefficients except the intercept are jointly zero is rejected at the 1% critical level, implying that the model has good explanatory power.

The coefficients on each attribute indicate the impact of that attribute on the probability of choosing the hypothetical asthma alternative. For example, cost has a negative impact on the decision to choose the hypothetical asthma medication. This is in accord with the property of the IUF outlined in Section 4. All effects were in the expected direction. The effect of 6 levels of attributes were statistically significant at the 1 percent level, 1 at the 5 percent level and 3 at the 10 percent level. The alternative specific constant is included to indicate a general preference for either the hypothetical scenario or for the respondents' current asthma medication. A preference for the former would be represented by a positive coefficient while a preference for the latter would be indicated by a negative coefficient. Given the ASC in the above model is negative this indicates that respondents have a general preference for their current asthma medication.

Welfare Calculations

The parameter estimates from the random utility model estimated above are employed to calculate the welfare measures associated with changes in the described asthma medication. The process is as follows: the parameters estimated above, the βs , are used in conjunction with the attribute levels, the Xs , to calculate numerical values for the conditional IUFs. The values of the IUFs, V_i, are then used in equation (4.17) to indirectly calculate the compensating variation.

Using this process, moving from the worst to the best hypothetical asthma medication results in a CV of $34.72 per prescription. This figure was calculated by changing all attributes of the hypothetical drug from the worst to best levels. This implies that if the quality of all the characteristics of the hypothetical drug were to improve, the respondent would have to have $34.72 taken away from them to leave them at the same utility level as they were prior to the quality improvement. That is, the respondent would be willing to pay $34.72 to secure the improved asthma medication.

---

high speed rail choices was dropped from the analysis of travel choice because that alternative was only chosen by 2 people.

These results are preliminary and drawn from SP data collected from a small sample and as such, should be interpreted as demonstrating the method of deriving welfare measures from SPDCM data that are consistent with microeconomic welfare theory rather than as a definitive result.

7. CONCLUSION AND AREAS FOR FUTURE RESEARCH

An essential component of policy evaluation in the health sector is the measurement of changes in welfare arising from new policies or modifications to existing policy. SPDCM offers one source of data from which to value changes in welfare. The discrete choice model used in SPDCM focuses on a comparison of the alternatives in the choice set and their underlying characteristics. SPDCM is therefore well suited to estimating changes in welfare from problems characterised by substitution among alternatives as is often the case in the health sector. SPDCM can be used as an alternative to RP data or more importantly can be used in cases where RP data is lacking.

This paper has investigated the use of SPDCM data as a source of information from which to indirectly derive measures of welfare that are consistent with microeconomic welfare theory. To do this, an overview of SPDCM was first provided, including its uses, underlying theory, methods and application. The theory of exact welfare measurement was reviewed in the familiar case of continuous data followed by a discussion of how these methods have been modified in order to be used with SPDCM data, that is, to take account of discrete data within a random utility context.

A review of the health economics literature highlighted that the methods used to derive estimates of WTP from SPDCM to date differ from the theoretically correct method suggested by Small and Rosen. A key difference is that the few studies that have estimated WTP figures from SPDCM data in the health economics field use a method that does not take account of the probability of choosing each alternative in the choice set. As such, a change in the levels of the attributes could potentially reduce the probability that the good or service in question will be chosen as respondents have the opportunity to substitute to more attractive alternatives in the choice set. The method used in the literature to date is appropriate, however, when the chosen alternative is known with certainty. Alternatively, it can be used as an approximation to the change in welfare if the change in the quality of the attributes is very small.

The feasibility of the theoretically consistent method proposed by Small and Rosen was demonstrated in an application to elicit the CV for asthma medication. This represents the first application of this method using SPDCM data in the health sector. The results of this application are purely preliminary. Due to the small sample, no attempt should be made to generalise the results from this sample to the population of asthmatics. The results should be interpreted as demonstrating the method of deriving welfare measures from SPDCM data that are consistent with microeconomic welfare theory rather than as a definitive result.

There are a number of key questions which need to be addressed in future research before the method set out in this paper can be considered a standard tool. An important area for future research is around the validity of welfare measures calculated from SP data. In particular, how does what people say they would do compare to the actual decisions they make in the market place. This in turn relates to the validity of the SPDCM method in general. An important part of future research must be to test the assumption underlying SPDCM that the choices people make in hypothetical scenarios reflect their preferences. The general question of the validity and reliability of SPDCM experiments has been addressed elsewhere (Louviere, Hensher et al. 2000; Louviere 1988), however further investigation will be required in the health economics field with a particular focus on welfare measures estimated from this technique. One method of validation is to compare welfare measures derived from RP data with those derived from SPDCM data using the same group of individuals.

An appropriate extension of this research is to include income in the empirical application rather than approximating this with the negative of the price coefficient. Using each individual's income also allows for a separate WTP calculation per person which can be aggregated to produce a mean WTP over the study population to produce an overall welfare gain. Future work will allow marginal utility of income to vary which is likely to be more consistent with economic theory.

An important extension of this work is to calculate the WTP for a product or program using both the method currently used in the literature and the theoretically consistent method outlined in this paper in order to ascertain the implications of using the former rather than the latter.

Other areas for future research include: explicitly incorporating uncertainty into the attributes in the SPDCM experiment in order to derive welfare measures under uncertainty; investigating the use of non linear functional forms in the estimation of the conditional logit model and their impact on the resulting welfare measures; comparing welfare measures derived from SPDCM data with those derived via the contingent valuation method; and investigating whether preferences elicited from SPDCM methods satisfy the axioms of completeness, reflexivity, transitivity, convexity and monotonicity.

Once a number of the issues outlined above have been addressed, particularly the issue of validity, the method presented here to derive welfare measures from SPDCM data could potentially move from being theoretically plausible to become a practical tool in the health economics field and potentially fill the gap left by the relative lack of RP data in the health sector.

APPENDIX ONE

SPDCM Literature Search

Table 1 Electronic databases accessed for literature search

Database	Period Covered
Medline	1980-2001
Premedline	2001
Econlit	1980-2001
Embase	1980-2001

Table 2 Search terms used for literature search

SPDCM	Health	WTP
Conjoint analysis		Welfare
Discrete Choice Experiments		Compensating Variation
Choice Modelling		Equivalent Variation

Inclusion and Exclusion criteria

All articles published in languages other than English were excluded as were articles published before 1980. Duplicates were excluded when more than one database was employed simultaneously.

Search Results

First, a search of the literature was carried out on three databases concurrently – Medline, Premedline, Econlit and Embase. The search terms “discrete choice”, “choice modeling”, “choice modelling”, “conjoint analysis” were used in conjunction with “health” – this yielded 86 citations eligible for inclusion.

The second stage of the search was to determine the number of articles within the above set that that also contained the key words or subject headings “willingness to pay”, “welfare”, “compensating variation”, or “equivalent variation”. To do this, these key words were combined with the above results. A total of 26 of the 86 original citations contained the relevant key words. On obtaining the articles they were screened to ascertain their relevance from which the number of relevant studies was further reduced to 15. The literature review was supplemented by a review of references from bibliographies of relevant papers.

APPENDIX TWO

Example of a hypothetical scenario from the asthma SPDCM questionnaire

Total cost of the drug to you for a 3 month supply is:	*Free*
In order to pick up your prescription and repeats for each 6 month period you will have to go to the chemist:	*Twice*
You take the drug via:	*a tablet*
You need to take the drug:	*Once a day*
On this drug you will experience:	*Exercise breathlessness, cough or wheeze once a week*
Your doctor says that this drug:	*is the best for your asthma*
Compared to most people in your age group you are able to participate in:	*all usual daily activities without the use of a reliever*
You are able to participate in:	*all the sporting or strenuous activity you want without difficulty*
You measure your morning peak flow:	*Never*
This drug will cause:	*no side effects*

Suppose you had the option of taking the medication described in the situation above. Considering how severe your asthma is, which option would you prefer out of:

☐ Your current preventer medication

☐ The preventer medication in the situation above

☐ No preventer medication

REFERENCES:

Adamowicz, W., Louviere, J. and Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Management and Economics,* 26: 271-292.

Adamowicz, W., Swait, J., Boxall, P., Louviere, J. and Williams, M. (1997). Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental and Management,* 32: 65-84.

Adamowicz, W. L., Boxall, P. C., Louviere, J. J., Swait, J. and Williams, M. (1999). Stated-Preference Methods for Valuing Environmental Amenities. Valuing Environmental Preferences. Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries. I. J. Bateman and Willis, K. G. Oxford, Oxford University Press: 460-479.

Ben - Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Applications to Travel Demand*. Cambridge, The MIT Press.

Bennett, J. and Blamey, R. (2001). The strengths and weaknesses of enviornmental choice modelling. The choice modelling approach to environmental valuation. J. Bennett and Blamey, R. Cheltonham, Edward Elgar Publishing Limited: 227-242.

Bockstael, N. E., McConnell, K. E. and Strand, I. E. (1991). Recreation. Measuring the Demand for Environmental Quality. J. B. Braden and Kolstad, C. D. Amsterdam, North Holland: 227-270.

Braden, J. B. and Kolstad, C. D. (1991). Summary and Conclusions. Measuring the Demand for Environmental Quality. J. B. Braden and Kolstad, C. D. North Holland, Elsevier: 323-332.

Chipman, J. S. and Moore, J. C. (1992). Compensating variation as a measure of welfare change. Economic theory and international trade : essays in memoriam, J. Trout Rader. W. Neuefeind and Riezman, R. G. New York, Springer-Verlag: 93-128.

Clarke, P. M. (1998). Cost-benefit analysis and mammographic screening: A travel cost approach. *Journal of Health Economics,* 17: 767-787.

Deaton, A. and Meullbauer, J. (1980). *Economics and Consumer Behaviour*. Cambridge, Cambridge University Press.

Feldman, R. (1994). The Welfare Economics of a Health Plan Merger. *Journal of Regulatory Economics,* 6: 67-86.

Gertler, P. (1987). Are User Fees Regressive? The Welfare Implications of Health Care Financing Proposals in Peru. *Journal of Econometrics,* 36: 67-88.

Griliches, Z. (1961). Hedonic price indexes for automobiles: An econometric analysis of quality change. The price statistics of Federal Government, NBER**:** 173-196.

Gyrd-Hansen, D. (2000). Cost benefit analysis of mammography screening in Denmark based on discrete ranking data. *International Journal of Technology in Health Care,* 16: 811-821.

Hanemann, W. M. (1982). *Applied welfare analysis with qualitative response models*. Berkeley, Publisher.

Hanemann, W. M. (1991). Willingness to Pay and Willingness to Accept: How Much Can They Differ? *American Economic Review,* 81: 635-47.

Hanemann, W. M. (1996). *Welfare analysis with discrete choice models*. Berkekey, Publisher.

Hanemann, W. M. (1999). The Economic Theory of WTP and WTA. Valuing Environmental Preferences. Theory and Practice of the Contingent Valuation Method in the US, EU and Developing Countries. I. J. Batemann and WIllis, K. G. Oxford, Oxford University Press**:** 42-96.

Hicks, J. R. (1939). The foundations of welfare economics. *The Economic Journal,* 49: 696-712.

Hicks, J. R. (1942). Consumers' surplus and index numbers. *Review of Economic Studies,* 9: 126-137.

Johnson, F. R., Desvousges, W. H., Ruby, M. C., Stieb, D. and De Civita, P. (1998). Eliciting stated health preferences: an application to willingness to pay for longevity. *Medical Decision Making,* 18: S57-67.

Johnson, F. R., Mathews, K. E. and Bingham, M. F. (2000). *Evaluationg welfare-theoretic consistency in multiple response stated preference surveys*. Durham, Triangle Economic Research.

King, M. A. (1983). Welfare Analysis of Tax Reforms Using Household Data. *Journal of Public Economics,* 21: 183-214.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy,* 74: 132-57.

Lancaster, K. (1991). *Modern Consumer Theory*. Vermont, Edward Elgar Publishing.

Louviere, J. (1988). Conjoint analysis modelling of stated preferences: a review of theory methods recent developments and external validity. *Journal of Transport Economics and Policy,* 22: 93-119.

Louviere, J. J., Hensher, D. A. and Swait, J. D. (2000). *Stated Choice Methods Analysis and Application*. Cambridge, Cambridge University Press.

Maddala, G. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge, Cambridge University Press.

Mas-Colell, A., Whinston, M. D. and Green, J. R. (1995). *Microeconomic Theory*. Oxford, Oxford University Press.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. Frontiers of econometrics. P. Zarembka. New York, Academic Press**:** 105-142.

McFadden, D. (1997). *Measuring willingness to pay for transportation improvements*. , Publisher.

McFadden, D. (1999). Computing willingness to pay in random utility models. Trade, theory and econometrics: essays in the honour of John S Chipman. J. S. Chipman, Moore, J. and Riezman, R. London, Routledge.

McKenzie, L. (1957). Demand theory without a utility index. *The Review of Economic Studies,* 24: 185-189.

Puig-Junoy, J. (1998). Why Do Patients Prefer Hospital Emergency Visits? A Nested Multinomial Logit Analysis for Patient-Initiated Contacts. *Health Care Management Science,* 1: 39-52.

Ryan, M. (1999). A role for conjoint analysis in technology assessment in health care? *International Journal of Technology Assessment in Health Care,* 15: 443-57.

Ryan, M. (1999). Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. *Social Science and Medicine,* 48: 535-546.

Ryan, M. (2000). Sensitivity of Willingness to Pay Estimates to the Level of Attributes in Discrete Choice Experiments. *Scottish Journal of Political Economy,* 47: 504-24.

Ryan, M. and Huges, J. (1997). Using conjoint analysis to assess women's preferences for miscarriage management. *Health Economics,* 6: 261-273.

Samuelson, P. A. (1974). Complementarity-An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory. *Journal of Economic Literature,* 12: 1255-89.

Scott, A. (2001). Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics. *Journal of Health Economics,* 20: 329-347.

Slesnick, D. T. (1998). Empirical approaches to the measurement of welfare. *Journal of Economic Literature,* 36: 2108-2165.

Small, K. A. and Rosen, H. S. (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica,* 49: 105 - 130.

Sugden, R. and Williams, A. (1978). *The Principles of Practical Cost-Benefit Analysis*. Oxford, Oxford University Press.

Trajtenberg, M. (1989). The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners. *Journal of Political Economy,* 97: 444-79.

Varian, H. R. (1992). *Microeconomic Analysis*. New York, Norton.

Willig, R. D. (1976). Consumer's Surplus without Apology. *American Economic Review,* 66: 589-97.

Wriglesworth, J. L. (1987). The Three Consumer Surpluses as Individual Welfare Measures. *Scottish Journal of Political Economy,* 34: 230-48.

UTS:MCU / JOB 10386 / JANUARY 2005 / ONLINE / UTS CRICOS PROVIDER CODE: 00099F